A peer-reviewed open-access journal
NeoBiota
Advancing research on alien species and biological invasions

NeoBiota 98: 319–334 (2025)
DOI: 10.3897/neobiota.98.147310

Research Article

A method to quantify jump dispersal of invasive species from occurrence data: the case of the spotted lanternfly, *Lycorma delicatula*

Nadège Belouard[1,2], Sebastiano De Bona[1,3], Matthew R. Helmus[1], Isabella G. Smith[1], Jocelyn E. Behm[1]

1 *Integrative Ecology Lab, Center for Biodiversity, Department of Biology, Temple University, 1925 N. 12th Street, Philadelphia, PA 19122, USA*
2 *ECOBIO (Ecosystèmes, Biodiversité, Evolution), Univ Rennes, CNRS, Rennes, France*
3 *Department of Biology, Bryn Mawr College, Bryn Mawr, Pennsylvania, USA*

Corresponding author: Nadège Belouard (nadege.belouard@gmail.com)

Abstract

The accuracy of predicting the spread of biological invasions is improved if models explicitly incorporate the two main dispersal mechanisms: diffusive spread and jump dispersal. However, quantitative methods for differentiating these two mechanisms in spatial occurrence data are lacking. We present jumpID, an R package using directional analysis of occurrence data to distinguish between jump dispersal and diffusive spread in biological invasions. We applied this method to occurrence data from the spotted lanternfly (*Lycorma delicatula*) invasion in the US, a pest rapidly expanding its range and impacting the forest and grape industries. Application of jumpID to a dataset of 123,542 occurrence records of spotted lanternfly uncovered 152 dispersal jumps between 2014–2022, with the first jump in 2017, three years after spotted lanternfly's first find. More than half of the dispersal jumps started satellite invasions the year after their detection. The average jump distance did not change over time, with 89% of jumps shorter than 200 km and just three jumps farther than 300 km. The overall spread rate was 41 ± 24 SD km/year, but reduced to 25 ± 11 SD km/year when considering diffusive spread only. Estimating jump dispersal enhances our understanding of species' dispersal mechanisms, provides more robust estimates of diffusion rates for spread models, and helps determine the perimeters for containment and control measures. The R package jumpID is openly available to facilitate invasive spread analysis. jumpID equips scientists and managers with a tool to separate the spread of invasive species into diffusion and jump dispersal components, allowing for more precise parameterization of spread models and directly informing management strategies. Application of jumpID to the spotted lanternfly system indicates that management efforts targeting jump dispersal should focus on a 200-km buffer around the invasion boundary which is where 89% of jumps occur.

Key words: Human-assisted dispersal, invasion boundary, invasion front, invasion spread, invasive species management, long-distance dispersal, secondary diffusion, stratified dispersal

OPEN ACCESS

Academic editor: Deepa Pureswaran
Received: 21 January 2025
Accepted: 25 April 2025
Published: 19 May 2025

Citation: Belouard N, De Bona S, Helmus MR, Smith IG, Behm JE (2025) A method to quantify jump dispersal of invasive species from occurrence data: the case of the spotted lanternfly, *Lycorma delicatula*. NeoBiota 98: 319–334. https://doi.org/10.3897/neobiota.98.147310

Copyright: © Nadège Belouard et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

Biological invasions are a costly component of global change and a threat to biodiversity and natural resources (Simberloff et al. 2013; Diagne et al. 2021). Predicting the establishment and spread of invasive populations is crucial to developing control strategies and limiting their impacts (e.g. Barbet-Massin et

al. 2018; Verdasca et al. 2021; van Rees et al. 2022). Projections of potential distributions of invasive species have recently started to incorporate dispersal processes to improve the accuracy of the predictions (Václavík and Meentemeyer 2009; Briscoe et al. 2019; Srivastava et al. 2021). In this context, dispersal is often accounted for as a diffusion process through the short-distance, continuous colonization of adjacent locations (Wilson et al. 2009; Lockwood et al. 2013). However, most species expand by stratified dispersal, a combination of short- and long-distance dispersal (Shigesada et al. 1995; Ciosi et al. 2011; Takahashi and Park 2020), and it is widely recognized that rare long-distance dispersal events (LDD) have major consequences for species spread (Trakhtenbrot et al. 2005; Hallatschek and Fisher 2014; Wu et al. 2023). In particular, extra-range LDD result in the establishment of satellite populations at locations away from the invasion range, beyond the limits of the continuous spread, that appear as discontinuous "jumps" in the spatial distribution of an invasive population (Suarez et al. 2001; Nathan 2006; Botella et al. 2022). These dispersal jumps, often due to human-assisted dispersal, become starting points for satellite invasions, a major concern for previously uninvaded areas (e.g. Lanner et al. 2020; Eyer et al. 2021). Yet, the stochastic nature of jump dispersal makes it difficult to identify, predict and incorporate into the projections of potential distributions of invasive species (Nathan et al. 2003; Tobin and Robinet 2022; Wu et al. 2023).

Because of the rarity of LDD, quantifying past dispersal jumps is the best way to understand the distance, frequency, and location at which dispersal jumps occur (Suarez et al. 2001; Nathan et al. 2003; Chivers and Leung 2012). Previously, jump dispersal has been identified with case-specific methods, including visual analyses of distribution data, theoretical models, genetics, or simply expert opinion (e.g. Bossenbroek et al. 2001; Signorile et al. 2016; Eyer et al. 2021; Lee et al. 2022). With increasingly extensive, coordinated survey efforts providing a large amount of occurrence data for invasive species (e.g. Verdasca et al. 2021; Dart et al. 2022; De Bona et al. 2023), there is an opportunity to leverage comprehensive datasets to develop generalized, quantitative methods for identifying dispersal jumps across systems.

Methods using occurrence data for distinguishing diffusive spread from jump dispersal in an ongoing invasion must account for three characteristics of invasions. First, the spread of biological invasions is often anisotropic, meaning diffusion occurs at different rates in different directions from the origin, which leads to invasion ranges with irregular shapes. Dispersal jumps in a direction with slow diffusive spread may be closer to the introduction site than parts of the diffusive range associated with faster spread. Detecting dispersal jumps with occurrence data thus requires delineation of the limits of the diffusive spread in different directions, or else dispersal jumps will be missed in slower progressing sectors. Second, secondary diffusion, i.e. the continuous spread starting from a dispersal jump, constitutes new continuous spread fronts around dispersal jumps (Shigesada et al. 1995; Suarez et al. 2001) that must not be mistaken for new dispersal jumps. Third, methods must account for a temporal component and detect the progression of the continuous spread front over time in both the initial diffusion range and around dispersal jumps to detect jumps only outside of previously colonized areas.

Here, we describe jumpID, a generalized method for distinguishing dispersal jumps from diffusive spread in biological invasions. jumpID identifies dispersal jumps based on spatial occurrence data using a conservative directional analysis to describe

the general patterns of jump dispersal and diffusive spread based on occurrence data. jumpID offers a flexible solution for scientists and stakeholders to understand dispersal in invasive populations, facilitating analyses of the features associated with spread patterns, and predictions about areas at high risk of future dispersal jumps.

As a case study, we applied jumpID to spatial and temporal occurrence data for the spotted lanternfly (*Lycorma delicatula*, hereafter SLF) planthopper (De Bona et al. 2023). Native to China and first discovered in the US in Pennsylvania in 2014, SLF is considered an invasive pest because of the threats it presents to the grape and forest industries (Urban and Leach 2023). SLF have established numerous satellite populations despite the enforcement of quarantine zones (see e.g. Pennsylvania Department of Agriculture 2021). Spread models suggest the existence of human-mediated jump dispersal (Ladin et al. 2023) and that SLF can easily spread via jumps to California in the next decade (Jones et al. 2022). SLF is thus an ideal species for demonstrating the utility of jumpID.

Methods

Aim and data requirements

The aim of jumpID is to distinguish dispersal jumps from diffusive spread in biological invasions based on spatial occurrence data. In jumpID, potential jumps are positive occurrences located beyond the limit of diffusive spread for the invasion. jumpID is freely available in R (R Core Team 2023) from an online GitHub repository containing tutorials at https://ieco-lab.github.io/jumpID.

jumpID requires at a minimum a dataset of presences that includes the geographic coordinates at which surveys were conducted. Additionally, datasets including absence data allow systematic verifications of survey effort in presumed uninvaded areas. Datasets with a temporal dimension also allow inquiries about temporal dynamics of spread as well as the identification of jumps enveloped by the progression of diffusive spread over time. If a temporal dimension is included, it is best to split the dataset into time steps that correspond to the temporal scale at which dispersal happens, such as each generation for short-lived species or each year for long-lived species.

Jump identification

Here we provide a brief overview of the jumpID methodology, with detailed descriptions of each component to follow. jumpID detects dispersal jumps by reconstructing the progression of a biological invasion using a directional analysis of occurrence survey data. Starting from a reference point within the invaded range and moving outwards, jumpID identifies spatial discontinuities in the distribution of species' occurrences between the limit of the continuous, diffusive spread and outlying occurrences. All positive records past this discontinuity are potential jumps. This process is repeated for each time step to follow the progression of the continuous spread front. Potential jumps are then confirmed or discarded based on their distance to previous jumps or to secondary diffusion around those. Eventually, the initial occurrence data is divided into three categories: "diffusion", "dispersal jumps", and "secondary diffusion" around dispersal jumps (terms in italics defined in Table 1). A step-by-step tutorial is included in the jumpID GitHub repository.

Table 1. Definitions of frequently used terms in jumpID.

Term	Definition
Types of dispersal	
Diffusion	Continuous colonization of adjacent locations, also called short-distance dispersal or diffusive spread
Jump dispersal	Discontinuous spread of a species' spatial distribution, also called long-distance dispersal
Secondary diffusion	Continuous colonization of locations adjacent to a dispersal event
jumpID terms	
Reference point	Focal point from which all distances for each presence point in the occurrence dataset are calculated.
Distance to the reference point (refDist)	Distance between each presence point and the reference point.
sector	The invaded range is divided into smaller sections (sectors) centered on the reference point and potential jumps are identified within each sector separately.
Minimum jump distance (MJD)	The extent of the spatial discontinuity that makes a presence point a potential jump
Threshold point	The presence point with the greatest refDist that defines the boundary of diffusive spread.
Jump cluster	For temporal jumpID analyses, a group of jumps from the same time step that are all less than the MJD from at least one other jump within the cluster.
Rarefied dataset	Dataset of potential jumps that represents jump clusters as a single point (the jump closest to the centroid of the jump cluster) instead of including all individual jumps.

Data preparation

To start, a "reference point" is chosen for the dataset, and the geographic distance from the reference point to each presence point is calculated ("refDist", Fig. 1A). Ideally, the location of the initial introduction site should be used for the reference point (Renault 2020), or if that is unknown, the center of the invasion core can be used.

To account for anisotropy in species spread, jumpID divides the invaded range into "sectors" centered on the reference point and identify potential jumps within each sector separately (Fig. 1B). Dividing space in any number of sectors will produce accurate results, as the list of potential jumps by sector is then screened through additional pairwise distance calculations across sectors. By default, jumpID divides space into 8 sectors. A demonstration of how computation time may vary with the number of sectors is provided in the jumpID GitHub repository.

Identification of thresholds, jumps, and secondary diffusion

We define the "minimum jump distance "(MJD) as the extent of the spatial discontinuity that makes a presence point a potential jump for a given case study. As outlined by Nathan et al. (2003), the choice of a MJD is inherently case-specific. It must correspond to a high percentage of their dispersal distance that includes all known dispersal mechanisms for outlying points to qualify as dispersal jumps (Trakhtenbrot et al. 2005). We suggest searching the literature for this range of dispersal distances in the invasion of interest, in the species' native range, or in the closest available species. The density of presence data will determine the lowest MJD that can be set in the study system, because the MJD must be higher than the distance separating positive records from continuous spread for jumpID to accurately detect the invasion front. In early invasions with sparse data, the number of sectors may be decreased to provide more data points for the directional analysis. Additionally, the MJD may be adjusted depending on the purpose of the jump analysis. For example, national stakeholders may only be interested in the characteristics of large jumps leading to the colonization of other states or countries (> 100 km), while scientists or local stakeholders may rather be interested in acquiring data on all dispersal jumps to fine tune spread models.

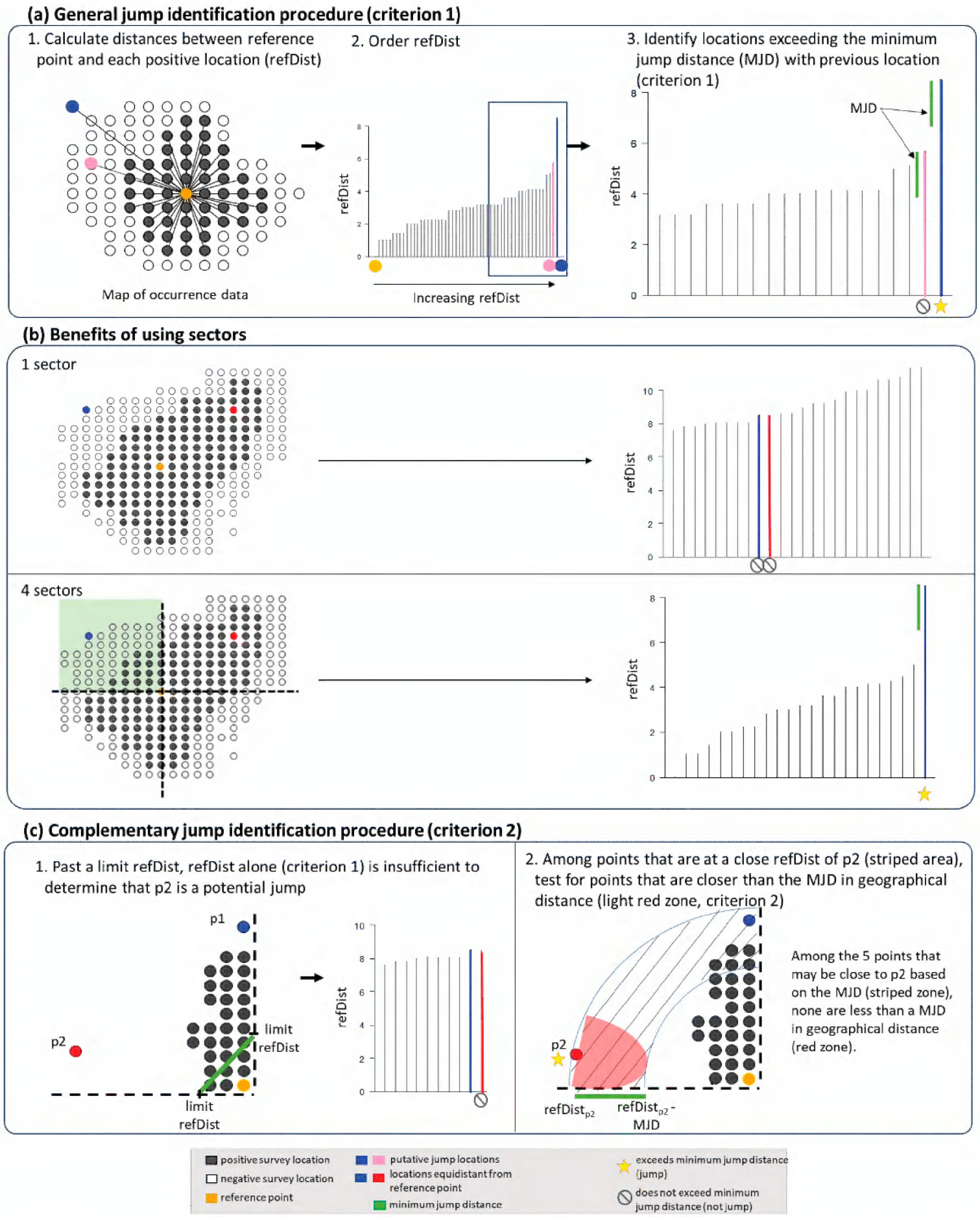
(a) General jump identification procedure (criterion 1)
1. Calculate distances between reference point and each positive location (refDist)
2. Order refDist
3. Identify locations exceeding the minimum jump distance (MJD) with previous location (criterion 1)
Map of occurrence data
refDist
Increasing refDist
MJD
(b) Benefits of using sectors
1 sector
4 sectors
(c) Complementary jump identification procedure (criterion 2)
1. Past a limit refDist, refDist alone (criterion 1) is insufficient to determine that p2 is a potential jump
p1
p2
limit refDist
2. Among points that are at a close refDist of p2 (striped area), test for points that are closer than the MJD in geographical distance (light red zone, criterion 2)
Among the 5 points that may be close to p2 based on the MJD (striped zone), none are less than a MJD in geographical distance (red zone).
refDist$_{p2}$
refDist$_{p2}$ - MJD
positive survey location
negative survey location
reference point
putative jump locations
locations equidistant from reference point
minimum jump distance
exceeds minimum jump distance (jump)
does not exceed minimum jump distance (not jump)

Figure 1. Graphical representation of the workflow used in jumpID **A** representation of the steps leading to jump identifications using the minimum jump distance (criterion 1). The blue point is identified as a jump, while the pink point is not because the difference in distances between the pink point and the preceding point is less than the minimum jump distance (green) **B** graphical demonstration of how a directional analysis accurately identifies dispersal jumps. Since the blue point and red point are equidistant from the reference point (yellow), with only one sector, the blue point is not recognized as a dispersal jump because of its location relative to the shape of the invasion. With 4 sectors, however, the blue point is correctly identified as a dispersal jump using the minimum jump distance (green) **C** graphical demonstration of how criterion 2 identifies potential jumps at longer distances from the reference point. Past a limit "refDist", points may be more distant between them than in "refDist", causing some jumps not to be detected using criterion 1 alone. In this case, a direct, complementary test of the presence of points within an inward MJD of a point (light red zone) determines whether the focal point is a potential jump.

In each sector, presence points are ordered by increasing "refDist" which allows for a directional analysis to detect spatial discontinuities in species presence that mark the limit between diffusive spread and potential jumps (Fig. 1B). Two consecutive presence points (p1, p2) whose difference in "refDist" is greater than the MJD indicate that p2 is a potential jump (criterion 1, Fig. 1A). Because of the gradual increase in sectors' widths over space, past a certain "refDist", two presence

points from the same sector may be separated by greater than the MJD from each other, but not their "refDist" (Fig. 1C). Thus, past this limit "refDist", the use of criterion 1 only is insufficient to detect all potential dispersal jumps, and jumpID evaluates a second criterion if criterion 1 does not detect a jump. For criterion 2, jumpID directly verifies that there is no point closer than a MJD in geographic distance inward of p2 among all points with a "refDist" within [$refDist_{p2}$ – MJD; $refDist_{p2}$]. If criterion 2 is verified, p2 is considered as a potential jump. Criterion 1 filters out most points and limits the longer pairwise distance analysis for criterion 2 to a subset of points that require this complementary test.

If a potential jump is detected according to criteria 1 or 2, the last point before the discontinuity (p1) defines the boundary of diffusive spread ("threshold point") and all remaining points after the discontinuity are listed as potential jumps, starting at p2. Optionally, the user may require the existence of absence points between the "threshold point" and a potential jump to attest that the observed discontinuity is not due to a lack of survey effort.

For datasets with a temporal component, we use the same approach separately for each time step, starting at the earliest time step. For each subsequent time step, we identify potential jumps using the cumulative data from the previous time steps.

Then, from the list of potential jumps from all sectors and all years, potential jumps that are located less than the MJD from any presence point from previous time steps are conservatively eliminated as jumps. In this case, if they are less than the MJD from any "threshold point", they are considered as diffusion; if they are less than the MJD from a previous jump point, they are considered secondary diffusion from this jump.

Overall, the jump analysis thus splits the original presence dataset into three separate datasets: diffusion (core invasion), dispersal jumps, and secondary diffusion from jumps.

Post-treatment of jump clusters

A "jump cluster" is a group of jumps from the same time step that are all less than the MJD from at least one other jump within the cluster. Jump clusters can result from repeated independent jumps to the location in the same time step, from a single jump that led to diffusive spread during the same time step, or even from a jump that diffused for several time steps before the location was surveyed. Because these clusters may not represent independent jumps, jumpID offers the possibility to rarefy the list of dispersal jumps to keep only the jump record that is the closest to the centroid of the jump cluster. The jump dataset obtained after this treatment is called a "rarefied dataset", as opposed to the full dataset that comprises all dispersal jumps, including clustered jumps. This optional rarefication treatment may be insightful depending on the ecological assumptions and the objectives of jump identification.

Case Study: Application of jumpID to SLF

We used a comprehensive open-access dataset summarizing SLF occurrence in the US between 2014 and 2022 at a 1-km^2 resolution (De Bona et al. 2023, data release v2_2023). The dataset comprised 123,542 sampled points, of which 48,601 contained established SLF populations (39%, Table 2, Suppl. material 1: fig. S1). These surveys showed that in just nine years, SLF established in 13 northeastern US states.

Table 2. Summary of jump records identified for each year.

Year	Jump records, annual (rarefied)	Jump records, cumulative (rarefied)
2014	0 (0)	0 (0)
2015	0 (0)	0 (0)
2016	0 (0)	0 (0)
2017	3 (2)	3 (2)
2018	58 (7)	61 (9)
2019	55 (13)	116 (22)
2020	79 (48)	195 (70)
2021	98 (43)	293 (113)
2022	94 (39)	387 (152)

Current evidence indicates that the entire US SLF invasion stems from a single introduction in Berks County, Pennsylvania (Barringer et al. 2015). To date, knowledge on SLF dispersal distances is minimal. Analyses of flight capabilities indicated that adults can only fly short distances (< 40 m) in a single bout (Baker et al. 2019; Myrick and Baker 2019; Wolfin et al. 2019). However, SLF can hitchhike on vehicle exteriors in both the mobile (Elsensohn et al. 2024) and egg (Urban 2020) stages and spread models have found dispersal jumps of more than 150 km in this species (Ladin et al. 2023). For the jumpID analysis, we defined a MJD of 15 km. This scale is likely much larger than their natural dispersal, yet it is small enough to obtain a comprehensive picture of jump dispersal in SLF, including relatively short dispersal jumps that may be caused by a different vector than statewide dispersal jumps. Since SLF complete their full life cycle in one year and have non-overlapping generations (Urban and Leach 2023), we split the dataset into time steps of one year.

We ran jumpID on this dataset with the introduction point as the reference point, a MJD of 15 km, and required checking for negative surveys in discontinuities. We rarefied the jump list with a MJD of 15 km. With the results obtained from jumpID, we calculated jump distances as the shortest distance between a diffusion point and each jump cluster to estimate the minimal distance that SLF had traveled to every jump location.

We conducted several statistical analyses to identify patterns in the jumpID output. We tested whether jump distances increased over the years using a linear model to check if more distant areas were at risk of colonization through jump dispersal over time. To estimate the difference in areas colonized by different dispersal mechanisms, we calculated the yearly invasion radius when considering the two spread patterns together or diffusion only, using the maximal "refDist" in each sector of the original dataset and of the diffusion dataset. To estimate the relative roles of diffusion and jump dispersal in accelerating the species progression, we tested whether the invasion radius increase was slower over time when considering diffusion only compared with all spread together using a linear model. The invasion radius in each sector was the response variable and it was log-transformed to meet assumptions of residuals normality. Year, spread type, and their interaction were the independent variables tested. The yearly spread was calculated as the yearly increase in the invasion radius for diffusion only. Finally, to determine the threat posed by jump dispersal, we determined if jumps were followed by secondary diffusion or rapidly enveloped by the core distribution by calculating if any diffusion point or secondary diffusion point were within the MJD of each jump after one year.

Results

A total of 387 presence records was associated with 152 independent jump events (rarefied jumps, Fig. 2). Clustered jumps were found in 59 of the 152 locations and the remaining 93 locations had only single jumps (Fig. 2). The largest three jump clusters comprised 30 (Winchester, VA, in 2018), 19 (Harrisburg, PA, 2018) and 18 points (Wilkes-Barre/Scranton area, PA, 2019). While jumps were found in all directions around the introduction site, they were disproportionately found west of the invasion range, across Pennsylvania, starting in 2019 (Fig. 2, Suppl. material 1: fig. S2).

We detected the first two dispersal jumps occurring in 2017, three years after SLF was first detected in the US (Figs 2, 3a). The number of new jumps at independent locations then increased every year to reach 48 in 2020, before decreasing to 43 in 2021 and 39 in 2022 (Fig. 3a, Table 2). By comparison, the number of new jumps (including jumps to the same location) rose from 3 to 58 between 2017 and 2018, then overall gradually increased to reach above 90 new yearly jumps in 2021 and 2022 (Fig. 3a, Table 2). The mean number of records per jump was 8.3 in 2018 and 4.2 in 2019, then consistently decreased to 1.6 in 2020, 2.3 in 2021, and 2.4 in 2022.

Jump distances did not increase over time (p = 0.71, Fig. 3b). The 152 independent jumps were located on average 87 ± 77 km (SD) from the closest continuous spread front, with 102 jumps (67%) of less than 100 km, 34 jumps (22%) of 100–199 km, 11 jumps (7%) of 200–299 km, 2 jumps (1%) of 300–399 km, and a single jump of over 400 km, in 2021 (Fig. 3c). After one year, 20% of jumps were enveloped by the progression of the diffusion, and among the other jumps, 53% were followed by secondary diffusion, whereas 47% had not spatially expanded (Suppl. material 1: fig. S3).

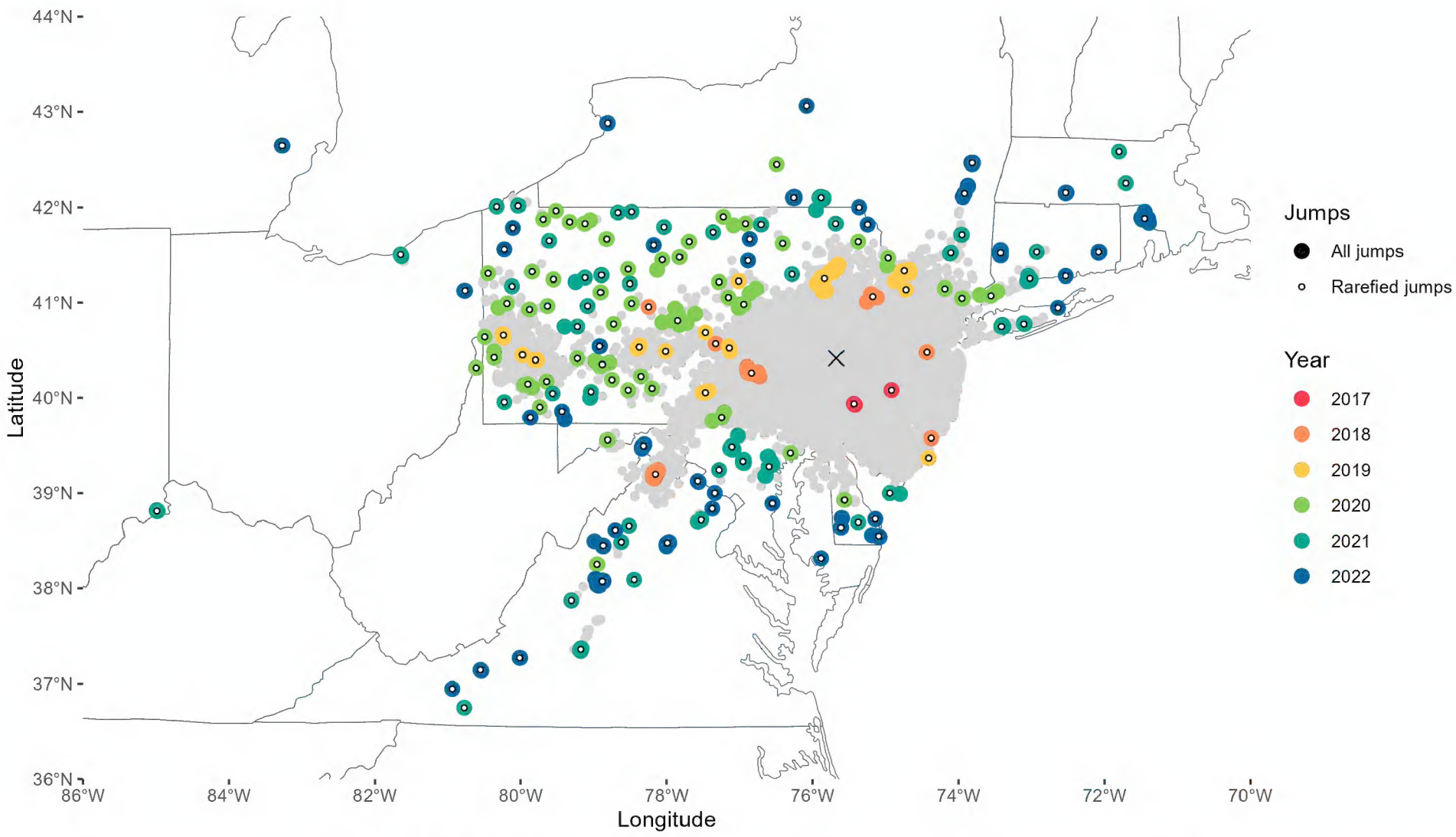

Figure 2. Map of the dispersal jumps identified based on spatial occurrence data of the invasive spotted lanternfly in the northeastern US, colored by year. Gray points represent all established populations recorded up to 2022, including secondary diffusion after jumps. The introduction site in Pennsylvania is signaled by a cross.

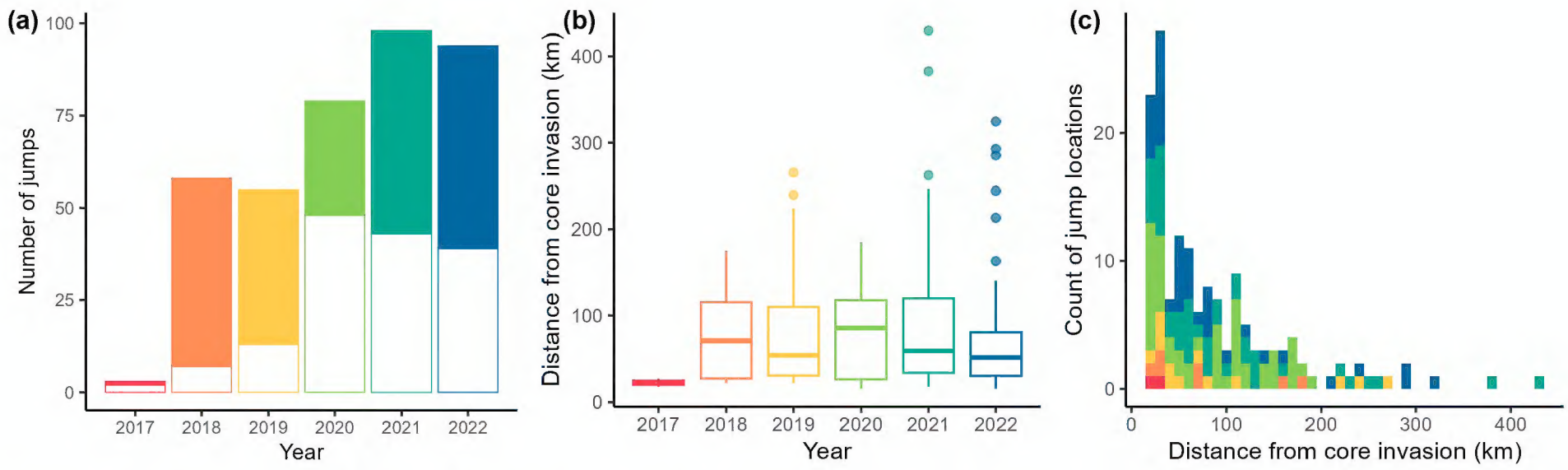

Figure 3. Characteristics of dispersal jumps identified based on spatial occurrence data of the invasive spotted lanternfly in the north-eastern US **a** number of jumps identified per year. The hollow part of bars represents the number of jump clusters (rarefied jumps), the entire bars represent the total number of jumps, including multiple jumps to a single location **b** distribution of the jump distance from the continuous spread front per year, based on rarefied jumps **c** dispersal kernel of jump distance, colored by year.

The invasion radius differed between years (p < 1e-15), spread types (p = 0.005) and their interaction (p = 0.023). In 2022, the invasion radius was on average 334 ± 204 km when including dispersal jumps and 204 ± 89 km when including only the core diffusion (Fig. 4). The radius of the core invasion (diffusive spread) increased on average by 25 ± 11 SD km per year, while the radius of the entire invasion range (diffusive spread, jumps, secondary diffusion) increased on average by 41 ± 24 SD km per year.

Discussion

Accounting for jump dispersal is critical for predicting the spread of invasive species. Jump dispersal allows satellite invasions to occur away from the core invasion, and thus management must enact control at multiple fronts (Suarez et al. 2001; Nathan et al. 2003). Even though jump dispersal is inherently stochastic due to the rarity of long-distance dispersal events that lead to the establishment of a population, risk zones can be identified based on past dispersal jumps. We presented jumpID, a quantitative method to differentiate dispersal jumps from diffusive spread in invasive species based on spatial occurrence data. We documented how this method provides quantitative information on the location, frequency and distance of dispersal jumps through the example of the spotted lanternfly (SLF, *Lycorma delicatula*) invasion in the US.

Analysis of dispersal jumps by jumpID

In our application of jumpID to the SLF dataset, 152 dispersal jumps were identified. The first dispersal jumps were detected in 2017, three years after the species was first discovered. SLF initially dispersed by diffusive spread and had established populations in a restricted area in 2014–2015. By 2017–2018, the invasion range had exploded in size due to stratified dispersal, the combination of diffusive spread and jump dispersal, because jump dispersal disproportionately increases the invasion radius (Suarez et al. 2001). Over half of dispersal jumps have been starting points for satellite invasions, highlighting the threat specifically posed by jump dispersal in the SLF invasion.

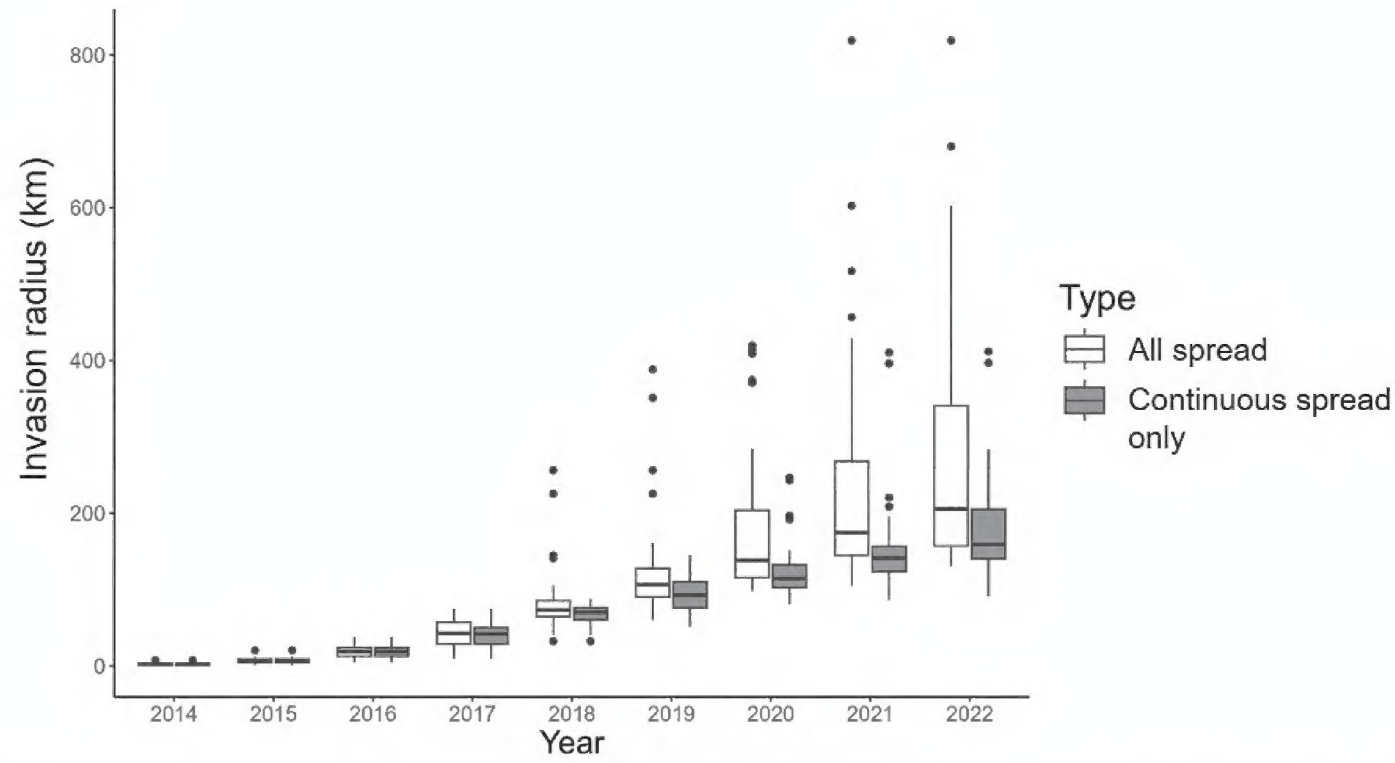

Figure 4. Change in the invasion radius depending on the spread patterns considered. Boxes represent the maximum radius of the invasion per sector for the entire dataset, jumps included (jumps and diffusion), and for diffusion only (diffusion only).

A single jump record was discovered at most jump locations. Such jumps may have included several individuals, or a single gravid female. Indeed, SLF are capable of multiple paternity, and a single mated female can lay multiple clutches of 30–100 eggs (Belouard and Behm 2023). A single gravid female or an egg mass may bring sufficient genetic diversity to sustain the establishment of a population following a dispersal jump. Jump clusters may be caused by a combination of massive propagule pressure (Lockwood et al. 2005) and vectors repeatedly transporting SLF to multiple proximate locations. However, it is most likely that the jump clusters identified were simply the delayed discovery of an original unique jump. In 2018, there was an average of 8 positive records at each jump location, and from 2020, it decreased to 1 or 2 positive records. This indicates that the preventative measures, the monitoring strategy, or both, were effective at decreasing the number of emigrants and/or discovering new infestations earlier, which are both known effective keys to contain biological invasions (Simberloff et al. 2013).

Even though the number of new jumps per year has been above 39 since 2020, it seems to be at least stabilizing in recent years compared to the trends from 2017–2020. Another encouraging signal is that jumps do not occur farther from the continuous spread front over time. It sets a theoretical upper boundary to the area at high risk of dispersal jumps ahead of the continuous spread front. The discovery of these patterns means that the distribution of jump distances may be directly incorporated into spread prediction models. A dual monitoring strategy may be applied ahead of the continuous spread front. First, monitoring may be systematic in a 200-km buffer around the continuous spread front, where most jumps (89%) happen. Second, identifying common characteristics of jump locations would result in a shortlist of at-risk locations to monitor outside the 200-km belt.

It may appear surprising that dispersal jumps were found mostly west of the invasion core. Indeed, jump dispersal is often due to human-assisted dispersal (Nathan et al. 2003) which biases dispersal in preferential axes associated with human activities (Bullock et al. 2018). As such, jumps to major cities along human transportation corridors (e.g., NYC, Baltimore, Washington DC) would have been expected to occur before jumps to central or western Pennsylvania. Therefore, SLF

may exploit specific transport infrastructures like freight rail lines that run east to west in this region rather than highways. Landscape analyses of dispersal jumps are a task for future studies to determine more attributes of jump dispersal in SLF based on the list established by jumpID, as well as determining how humans may be facilitating the spread of invaders in general (Bullock et al. 2018; Gippet et al. 2019).

Spread considerations and perspectives on jumpID

By separating dispersal jumps from diffusive spread and secondary diffusion, jumpID also allows a fine analysis of diffusion patterns. Basic estimates of spread rates showed a much slower progression of SLF over time after the jumpID analysis than when diffusion and jump dispersal are considered together. The major benefit of calculating estimates of both diffusive spread and jump dispersal for parameterizing spread models is more realistic spread predictions for invasive species (Robinet et al. 2017; Takahashi and Park 2020).

The choice of a minimal jump distance is a compromise between ecological knowledge, data density, and the purpose of the jump analysis. Indeed, a sparse occurrence dataset may constrain the user to adjust the MJD upwards of the value that would be set using ecological knowledge alone. Setting a MJD larger than the smallest true jump distances would result in unidentified short jumps, i.e. the jump distance kernel being truncated on the left. Intentionally overestimating the MJD may be desirable for users interested in the upper range of jump distances, such as dispersal events that make the species cross state borders. On the opposite, an underestimated MJD would most likely be encountered in systems with limited knowledge on dispersal distances combined with dense occurrence data. It would cause some presence points from the continuous spread to be identified as jumps. An underestimated MJD would be detected in summary statistics and on the maps provided by jumpID: a large proportion of presence data would be identified as jumps, the invasion front would progress very little every year, and most noticeably, jumps would be identified in regions that are visibly within the core invasion (e.g. on yearly maps such as Suppl. material 1: fig. S2). Finally, the MJD is somewhat independent of the continuous spread rate estimates. Diffusive spread may be partly human-assisted, and short-distance vectors accelerate the spread of invasive populations while keeping the invaded range continuous (Lockwood et al. 2013). Here, estimates of yearly spread by diffusion were larger than the minimal jump distance of 15 km considered in this analysis, showing that even short jumps can be identified ahead of highly dynamic invasion fronts. The minimal jump distance is thus not necessarily longer than the average or maximum yearly continuous spread rate.

jumpID is based on occurrence data only, and thus cannot determine the origin of a dispersal jump. For the same reason, jumpID is not intended to detect whether repetitive jumps occurred around the same location over time, as it cannot be differentiated from secondary diffusion around an initial jump. In all cases where a presence point could indicate either diffusive spread or jump dispersal, jumpID makes the conservative assumption that it is not a jump to favor independence in the data.

jumpID optionally accounts for absence data in the distribution of the invasive species as an indicator of survey effort within discontinuities in the species range. When this option is activated, at least one absence point after the continuous spread front is required to validate a dispersal jump. Although it is an imperfect measure of a species' true absence, this use of absence points in jumpID does not

alter the results of jumpID as in species distribution models (Václavík and Meentemeyer 2009; Chapman et al. 2019). When absence data are available, this criterion may be strengthened to request a higher number of absence data per gap.

Conclusion

Overall, jumpID is a flexible workflow that can be applied to any species with minimal ecological information to distinguish jump dispersal from diffusive spread based on spatial occurrence data. jumpID is freely available in a public repository and open to modifications to encourage scientists to adapt it to their species of interest. The jump identification and its simple, descriptive analysis is informative for summarizing areas of interest and the overall characteristics of jump dispersal, both temporally and spatially. It opens avenues for more in-depth analyses of jump dispersal and diffusive spread. The insights provided by jumpID results will help stakeholders focus survey efforts on locations predicted at highest risk of future dispersal jumps, thereby increasing the chances of early detection and local eradication before satellite invasions can take place (Simberloff 2014). jumpID thereby contributes to incorporating invasion science into management planning, as largely advocated by the scientific community (Diagne et al. 2020, van Rees et al. 2022).

Acknowledgments

We thank two anonymous reviewers for their comments on a previous version of this manuscript.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

This work was funded by the United States Department of Agriculture (USDA) Animal and Plant Health Inspection Service Plant Protection and Quarantine under agreements AP19PPQS&T00C251, AP20PPQS&T00C136, and AP22PPQS&T00C146; the USDA National Institute of Food and Agriculture Specialty Crop Research Initiative Coordinated Agricultural project award 2019-51181-30014; the USDA Tactical Sciences for Agricultural Biosecurity project award 2022-68013-37139; and the Pennsylvania Department of Agriculture under agreements C9400000036, C94000833, and C940001671.

Author contributions

NB conceived and performed the analysis and wrote the manuscript. JEB, SDB, MRH and IGS contributed critically to the analysis and edited the manuscript. All authors gave final approval for publication. Our study brings together authors from different countries, including scientists based in the country where the data were collected. All authors were engaged early on with the research and study design to ensure that the diverse sets of perspectives they represent were considered from the onset. In addition, local stakeholders were consulted, and their input was incorporated into the study design.

Author ORCIDs

Nadège Belouard https://orcid.org/0000-0002-7968-7735
Sebastiano De Bona https://orcid.org/0000-0001-6193-7073
Matthew R. Helmus https://orcid.org/0000-0003-3977-0507
Jocelyn E. Behm https://orcid.org/0000-0003-4220-4741

Data availability

jumpID is available at https://ieco-lab.github.io/jumpID, as well as all data and code used in this manuscript.

References

Baker TC, Smyers EC, Urban JM, Meng Z, Pagadala Damadaram KJ, Myrick AJ, Cooperband MF, Domingue MJ (2019) Progression of seasonal activities of adults of the spotted lanternfly, *Lycorma delicatula*, during the 2017 season of mass flight dispersal behavior in eastern Pennsylvania. Journal of Asia-Pacific Entomology 22(3): 705–713. https://doi.org/10.1016/j.aspen.2019.05.006

Barbet-Massin M, Rome Q, Villemant C, Courchamp F (2018) Can species distribution models really predict the expansion of invasive species? PLoS ONE 13: e0193085. https://doi.org/10.1371/journal.pone.0193085

Barringer LE, Donovall LR, Spichiger S-E, Lynch D, Henry D (2015) The first new world record of *Lycorma delicatula* (Insecta: Hemiptera: Fulgoridae). Entomological News 125(1): 20–23. https://doi.org/10.3157/021.125.0105

Belouard N, Behm JE (2023) Multiple paternity in the invasive spotted lanternfly (Hemiptera: Fulgoridae). Environmental Entomology 52(5): nvad083. https://doi.org/10.1093/ee/nvad083

Bossenbroek JM, Kraft CE, Nekola JC (2001) Prediction of long-distance dispersal using gravity models: Zebra mussel invasion of inland lakes. Ecological Applications 11(6): 1778–1788. https://doi.org/10.1890/1051-0761(2001)011[1778:POLDDU]2.0.CO;2

Botella C, Bonnet P, Hui C, Joly A, Richardson DM (2022) Dynamic species distribution modeling reveals the pivotal role of human-mediated long-distance dispersal in plant invasion. Biology (Basel) 11(9): 1293. https://doi.org/10.3390/biology11091293

Briscoe NJ, Elith J, Salguero-Gómez R, Lahoz-Monfort JJ, Camac JS, Giljohann KM, Holden MH, Hradsky BA, Kearney MR, McMahon SM, Phillips BL, Regan TJ, Rhodes JR, Vesk PA, Wintle BA, Yen JDL, Guillera-Arroita G (2019) Forecasting species range dynamics with process-explicit models: Matching methods to applications. Ecology Letters 22(11): 1940–1956. https://doi.org/10.1111/ele.13348

Bullock JM, Bonte D, Pufal G, da Silva Carvalho C, Chapman DS, García C, García D, Matthysen E, Delgado MM (2018) Human-mediated dispersal and the rewiring of spatial networks. Trends in Ecology & Evolution 33(12): 958–970. https://doi.org/10.1016/j.tree.2018.09.008

Chapman D, Pescott OL, Roy HE, Tanner R (2019) Improving species distribution models for invasive non-native species with biologically informed pseudo-absence selection. Journal of Biogeography 46(5): 1029–1040. https://doi.org/10.1111/jbi.13555

Chivers C, Leung B (2012) Predicting invasions: Alternative models of human-mediated dispersal and interactions between dispersal network structure and Allee effects. Journal of Applied Ecology 49(5): 1113–1123. https://doi.org/10.1111/j.1365-2664.2012.02183.x

Ciosi M, Miller NJ, Toepfer S, Estoup A, Guillemaud T (2011) Stratified dispersal and increasing genetic variation during the invasion of Central Europe by the western corn rootworm, *Diabrotica virgifera virgifera*. Evolutionary Applications 4(1): 54–70. https://doi.org/10.1111/j.1752-4571.2010.00133.x

Dart K, Latty T, Greenville A (2022) Citizen science reveals current distribution, predicted habitat suitability and resource requirements of the introduced African Carder Bee

Pseudoanthidium (*Immanthidium*) *repetitum* in Australia. Biological Invasions 24(6): 1827–1838. https://doi.org/10.1007/s10530-022-02753-2

De Bona S, Barringer L, Kurtz P, Losiewicz J, Parra GR, Helmus MR (2023) lydemapr: An R package to track the spread of the invasive spotted lanternfly (*Lycorma delicatula*, White 1845) (Hemiptera, Fulgoridae) in the United States. NeoBiota 86: 151–168. https://doi.org/10.3897/neobiota.86.101471

Diagne C, Leroy B, Gozlan RE, Vaissière A-C, Assailly C, Nuninger L, Roiz D, Jourdain F, Jarić I, Courchamp F (2020) InvaCost, a public database of the economic costs of biological invasions worldwide. Scientific Data 7(1): 277. https://doi.org/10.1038/s41597-020-00586-z

Diagne C, Leroy B, Vaissière A-C, Gozlan RE, Roiz D, Jarić I, Salles J-M, Bradshaw CJA, Courchamp F (2021) High and rising economic costs of biological invasions worldwide. Nature 592(7855): 571–576. https://doi.org/10.1038/s41586-021-03405-6

Elsensohn JE, Wolford S, Tabb A, Leskey T (2024) Experimental evidence supports the ability of spotted lanternfly to hitchhike on vehicle exteriors as a mechanism for anthropogenic dispersal. Royal Society Open Science 11(7): 240493. https://doi.org/10.1098/rsos.240493

Eyer P-A, Blumenfeld AJ, Johnson LNL, Perdereau E, Shults P, Wang S, Dedeine F, Dupont S, Bagnères A-G, Vargo EL (2021) Extensive human-mediated jump dispersal within and across the native and introduced ranges of the invasive termite *Reticulitermes flavipes*. Molecular Ecology 30(16): 3948–3964. https://doi.org/10.1111/mec.16022

Gippet JM, Liebhold AM, Fenn-Moltu G, Bertelsmeier C (2019) Human-mediated dispersal in insects. Current Opinion in Insect Science 35: 96–102. https://doi.org/10.1016/j.cois.2019.07.005

Hallatschek O, Fisher DS (2014) Acceleration of evolutionary spread by long-range dispersal. Proceedings of the National Academy of Sciences of the United States of America 111(46): E4911–E4919. https://doi.org/10.1073/pnas.1404663111

Jones C, Skrip MM, Seliger BJ, Jones S, Wakie T, Takeuchi Y, Petras V, Petrasova A, Meentemeyer RK (2022) Spotted lanternfly predicted to establish in California by 2033 without preventative management. Communications Biology 5(1): 1–9. https://doi.org/10.1038/s42003-022-03447-0

Ladin ZS, Eggen DA, Trammell TLE, D'Amico V (2023) Human-mediated dispersal drives the spread of the spotted lanternfly (*Lycorma delicatula*). Scientific Reports 13(1): 1098. https://doi.org/10.1038/s41598-022-25989-3

Lanner J, Huchler K, Pachinger B, Sedivy C, Meimberg H (2020) Dispersal patterns of an introduced wild bee, *Megachile sculpturalis* Smith, 1853 (Hymenoptera: Megachilidae) in European alpine countries. PLoS ONE 15: e0236042. https://doi.org/10.1371/journal.pone.0236042

Lee S-J, Cho H, Choi Y, Choi WI, Chung HI, Lim NO, Nam Y, Jeon S (2022) Path-finding algorithm as a dispersal assessment method for invasive species with human-vectored long-distance dispersal event. Diversity & Distributions 28(6): 1214–1226. https://doi.org/10.1111/ddi.13524

Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. Trends in Ecology & Evolution 20(5): 223–228. https://doi.org/10.1016/j.tree.2005.02.004

Lockwood JL, Hoopes MF, Marchetti MP (2013) Invasion ecology. Second edition. Wiley-Blackwell, Chichester, West Sussex, UK, 444 pp.

Myrick AJ, Baker TC (2019) Analysis of anemotactic flight tendencies of the spotted lanternfly (*Lycorma delicatula*) during the 2017 mass dispersal flights in Pennsylvania. Journal of Insect Behavior 32(1): 11–23. https://doi.org/10.1007/s10905-019-09708-x

Nathan R (2006) Long-distance dispersal of plants. Science 313(5788): 786–788. https://doi.org/10.1126/science.1124975

Nathan R, Perry G, Cronin JT, Strand AE, Cain ML (2003) Methods for estimating long-distance dispersal. Oikos 103(2): 261–273. https://doi.org/10.1034/j.1600-0706.2003.12146.x

Pennsylvania Department of Agriculture (2021) Spotted Lanternfly Quarantine Information. http://www.agriculture.pa.gov/Plants_Land_Water/PlantIndustry/Entomology/spotted_lanternfly/quarantine/Pages/default.aspx [January 26, 2018]

R Core Team (2023) R: A Language and Environment for Statistical Computing. https://www.R-project.org/

Renault D (2020) A review of the phenotypic traits associated with insect dispersal polymorphism, and experimental designs for sorting out resident and disperser phenotypes. Insects 11(4): 214. https://doi.org/10.3390/insects11040214

Robinet C, Suppo C, Darrouzet E (2017) Rapid spread of the invasive yellow-legged hornet in France: The role of human-mediated dispersal and the effects of control measures. Journal of Applied Ecology 54(1): 205–215. https://doi.org/10.1111/1365-2664.12724

Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. American Naturalist 146(2): 229–251. https://doi.org/10.1086/285796

Signorile AL, Lurz PWW, Wang J, Reuman DC, Carbone C (2016) Mixture or mosaic? Genetic patterns in UK grey squirrels support a human-mediated 'long-jump' invasion mechanism. Diversity & Distributions 22(5): 566–577. https://doi.org/10.1111/ddi.12424

Simberloff D (2014) Biological invasions: What's worth fighting and what can be won? Ecological Engineering 65: 112–121. https://doi.org/10.1016/j.ecoleng.2013.08.004

Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M, Pyšek P, Sousa R, Tabacchi E, Vilà M (2013) Impacts of biological invasions: What's what and the way forward. Trends in Ecology & Evolution 28(1): 58–66. https://doi.org/10.1016/j.tree.2012.07.013

Srivastava V, Roe AD, Keena MA, Hamelin RC, Griess VC (2021) Oh the places they'll go: Improving species distribution modelling for invasive forest pests in an uncertain world. Biological Invasions 23(1): 297–349. https://doi.org/10.1007/s10530-020-02372-9

Suarez AV, Holway DA, Case TJ (2001) Patterns of spread in biological invasions dominated by long-distance jump dispersal: Insights from Argentine ants. Proceedings of the National Academy of Sciences of the United States of America 98(3): 1095–1100. https://doi.org/10.1073/pnas.98.3.1095

Takahashi D, Park Y-S (2020) Spatial heterogeneities of human-mediated dispersal vectors accelerate the range expansion of invaders with source–destination-mediated dispersal. Scientific Reports 10(1): 21410. https://doi.org/10.1038/s41598-020-78633-3

Tobin PC, Robinet C (2022) Advances in understanding and predicting the spread of invading insect populations. Current Opinion in Insect Science 54: 100985. https://doi.org/10.1016/j.cois.2022.100985

Trakhtenbrot A, Nathan R, Perry G, Richardson DM (2005) The importance of long-distance dispersal in biodiversity conservation. Diversity & Distributions 11(2): 173–181. https://doi.org/10.1111/j.1366-9516.2005.00156.x

Urban JM (2020) Perspective: Shedding light on spotted lanternfly impacts in the USA. Pest Management Science 76(1): 10–17. https://doi.org/10.1002/ps.5619

Urban JM, Leach H (2023) Biology and management of the spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae), in the United States. Annual Review of Entomology 68. https://doi.org/10.1146/annurev-ento-120220-111140

Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling 220(23): 3248–3258. https://doi.org/10.1016/j.ecolmodel.2009.08.013

van Rees C, Hand B, Carter S, Bargeron C, Cline T, Daniel W, Ferrante J, Gaddis K, Hunter M, Jarnevich C, McGeoch M, Morisette J, Neilson M, Roy H, Rozance M, Sepulveda A, Wallace R, Whited D, Wilcox T, Luikart G (2022) A framework to integrate innovations in invasion science

for proactive management. Biological Reviews of the Cambridge Philosophical Society 97(4): 1712–1735. https://doi.org/10.1111/brv.12859

Verdasca MJ, Rebelo H, Carvalheiro LG, Rebelo R (2021) Invasive hornets on the road: Motorway-driven dispersal must be considered in management plans of *Vespa velutina*. NeoBiota 69: 177–198. https://doi.org/10.3897/neobiota.69.71352

Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in the way you move: Dispersal pathways affect invasion success. Trends in Ecology & Evolution 24(3): 136–144. https://doi.org/10.1016/j.tree.2008.10.007

Wolfin MS, Binyameen M, Wang Y, Urban JM, Roberts DC, Baker TC (2019) Flight dispersal capabilities of female spotted lanternflies (*Lycorma delicatula*) related to size and mating status. Journal of Insect Behavior 32(3): 188–200. https://doi.org/10.1007/s10905-019-09724-x

Wu Z-Y, Milne RI, Liu J, Nathan R, Corlett RT, Li D-Z (2023) The establishment of plants following long-distance dispersal. Trends in Ecology & Evolution 38(3): 289–300. https://doi.org/10.1016/j.tree.2022.11.003

Supplementary material 1

Additional information

Authors: Nadège Belouard, Sebastiano De Bona, Matthew R. Helmus, Isabella G. Smith, Jocelyn E. Behm

Data type: pdf

Explanation note: **fig. S1.** Overview of the data used. **fig. S2.** Results of jumpID, mapped per year. **fig. S3.** Map and occurrences of secondary diffusion per year. **table S1.** Summary of jump identification process per year.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.98.147310.suppl1